

1/7
SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANTS: James M. Anderson
Christina M. Van Itallie
- (ii) TITLE OF INVENTION: Human Occludin, Its Uses
and Enhancement of Drug Absorption Using Occlu-
din Inhibitors
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS
 - (A) ADDRESSEE: Yale University Medical School
Section of Pulmonary and Critical Care Medicine
Department of Internal Medicine
 - (B) STREET: 333 Cedar Street, LCI 105
 - (C) CITY: New Haven
 - (D) STATE: Connecticut
 - (E) COUNTRY: United States of America
 - (F) ZIP CODE: 065220-8057
- (v) COMPUTER READABLE FORM
 - (A) MEDIUM TYPE: 3.5" 1.44 Mb diskette
 - (B) COMPUTER: IBM PC
 - (C) OPERATING SYSTEM: MS DOS
 - (D) SOFTWARE: Word Processing
- (vi) CURRENT APPLICATION DATA
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: June 25, 2001
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA
 - (A) APPLICATION NUMBER: U.S. 09/142,732
 - (B) FILING DATE: September 15, 1998
 - (C) CLASSIFICATION: 530-350.000
- (viii) ATTORNEY INFORMATION
 - (A) NAME: Mary M. Krinsky
 - (B) REGISTRATION NO.: 32423
 - (C) REFERENCE/DOCKET NUMBER: OCR-754.CIP
- (ix) TELECOMMUNICATION INFORMATION
 - (A) TELEPHONE NUMBER: 203-773-9544
 - (B) TELEFAX NUMBER: 203-773-1183

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2312
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: cDNA

(v) FRAGMENT TYPE: complete sequence

(ix) FEATURE:

- (A) NAME/KEY: human occludin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCCTCTCTCC	ATCAGACACC	CCAAGGTTCC	ATCCGAAGCA	GGCGGAGCAC	50
CGAACGCACC	CCGGGGTGGT	CAGGGACCCC	CATCCGTGCT	GCCCCCTAGG	100
AGCCCGCGCC	TCTCCTCTGC	GCCCCGCCTC	TCGGGCCGCA	ACATCGCGCG	150
GTTCTTTTAA	CAGCGCGCTG	GCAGGGTGTG	GGAAGCAGGA	CCGCGTCCTC	200
CCGCCCCCTC	CCATCCGAGT	TTCAGGTGAA	TTGGTCACCG	AGGGAGGAGG	250
CCGACACACC	ACACCTACAC	TCCCGCGTCC	ACCTCTCCCT	CCCTGCTTCC	300
TCTTGCGGGA	GGCGGCAGGA	ACCGAGAGCC	AGGTCCAGAG	CGCCGAGGAG	350
CCGGTCTAGG	ACGCAGCAGA	TTGGTTTATC	TTGGAAGCTA	AAGGGCATTG	400
CTCATCCTGA	AGATCAGCTG	ACCATTGACA	ATCAGCCATG	TCATCCAGGC	450
CTCTTGAAAG	TCCACCTCCT	TACAGGCCTG	ATGAATTCAA	ACCGAATCAT	500
TATGCACCAA	GCAATGACAT	ATATGGTGGA	GAGATGCATG	TTCGACCAAT	550
GCTCTCTCAG	CCAGCCTACT	CTTTTTACCC	AGAAGATGAA	ATTCTTCACT	600
TCTACAAATG	GACCTCTCCT	CCAGGAGTGA	TTCGGATCCT	GTCTATGCTC	650
ATTATTGTGA	TGTGCATTGC	CATCTTTGCC	TGTGTGGCCT	CCACGCTTGC	700
CTGGGACAGA	GGCTATGGAA	CTTCCCTTTT	AGGAGGTAGT	GTAGGCTACC	750
CTTATGGAGG	AAGTGGCTTT	GGTAGCTACG	GAAGTGGCTA	TGGCTATGGC	800
TATGGTTATG	GCTATGGCTA	CGGAGGCTAT	ACAGACCCAA	GAGCAGCAAA	850
GGGCTTCATG	TTGGCCATGG	CTGCCTTTTG	TTTCATTGCC	GCGTTGGTGA	900
TCTTTGTTAC	CAGTGTTATA	AGATCTGAAA	TGTCCAGAAC	AAGAAGATAC	950
TACTTAAGTG	TGATAATAGT	GAGTGCTATC	CTGGGCATCA	TGGTGTTTAT	1000
TGCCACAATT	GTCTATATAA	TGGGAGTGAA	CCCAACTGCT	CAGTCTTCTG	1050
GATCTCTATA	TGGTTCACAA	ATATATGCCC	TCTGCAACCA	ATTTTATACA	1100
CCTGCAGCTA	CTGGACTCTA	CGTGGATCAG	TATTTGTATC	ACTACTGTGT	1150
TGTGGATCCC	CAGGAGGCCA	TTGCCATTGT	ACTGGGGTTC	ATGATTATTG	1200

TGGCTTTTGC TTTAATAATT TTCTTTGCTG TGAAAACCTCG AAGAAAGATG 1250
 GACAGGTATG ACAAGTCCAA TATTTTGTGG GACAAGGAAC ACATTTATGA 1300
 TGAGCAGCCC CCCAATGTCG AGGAGTGGGT TAAAAATGTG TCTGCAGGCA 1350
 CACAGGACGT GCCTTCACCC CCATCTGACT ATGTGGAAAG AGTTGACAGT 1400
 CCCATGGCAT ACTCTTCCAA TGGCAAAGTG AATGACAAGC GGTTTTATCC 1450
 AGAGTCTTCC TATAAATCCA CGCCGGTTCC TGAAGTGGTT CAGGAGCTTC 1500
 CATTAACTTC GCCTGTGGAT GACTTCAGGC AGCCTCGTTA CAGCAGCGGT 1550
 GGTAACTTTG AGACACCTTC AAAAAGAGCA CCTGCAAAGG GAAGAGCAGG 1600
 AAGGTCAAAG AGAACAGAGC AAGATCACTA TGAGACAGAC TACACAACCTG 1650
 GCGGCGAGTC CTGTGATGAG CTGGAGGAGG ACTGGATCAG GGAATATCCA 1700
 CCTATCACTT CAGATCAACA AAGACAACCTG TACAAGAGGA ATTTTGACAC 1750
 TGGCCTACAG GAATACAAGA GCTTACAATC AGAACTTGAT GAGATCAATA 1800
 AAGAACTCTC CCGTTTGGAT AAAGAATTGG ATGACTATAG AGAAGAAAGT 1850
 GAAGAGTACA TGGCTGCTGC TGATGAATAC AATAGACTGA AGCAAGTGAA 1900
 GGGATCTGCA GATTACAAAA GTAAGAAGAA TCATTGCAAG CAGTTAAAGA 1950
 GCAAATTGTC ACACATCAAG AAGATGGTTG GAGACTATGA TAGACAGAAA 2000
 ACATAGAAGG CTGATGCCAA GTTGTTTGAG AAATTAAGTA TCTGACATCT 2050
 CTGCAATCTT CTCAGAAGGC AAATGACTTT GGACCATAAC CCCGGAAGCC 2100
 AAACCTCTGT GAGCATCACA AAGTTTTTGGG TTGCTTTAAC ATCATCAGTA 2150
 TTGAAGCATT TTATAAATCG CTTTTGATAA TCAACTGGGC TGAACAACCTC 2200
 CAATTAAGGA TTTTATGCTT TAAACATTGG TTCTTGTATT AAGAATGAAA 2250
 TACTGTTTGA GGTTTTTAAAG CCTTAAAGGA AGGTTCTGGT GTGAACTAAA 2300
 CTTTCACACC CC 2312

(3) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 522

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: polypeptide

(v) FRAGMENT TYPE: complete sequence

(ix) FEATURE:

(A) NAME/KEY: human occludin

Met	Ser	Ser	Arg	Pro 5	Leu	Glu	Ser	Pro	Pro 10	Pro	Tyr	Arg	Pro	Asp 15
Glu	Phe	Lys	Pro	Asn 20	His	Tyr	Ala	Pro	Ser 25	Asn	Asp	Ile	Tyr	Gly 30
Gly	Glu	Met	His	Val 35	Arg	Pro	Met	Leu	Ser 40	Gln	Pro	Ala	Tyr	Ser 45
Phe	Tyr	Pro	Glu	Asp 50	Glu	Ile	Leu	His	Phe 55	Tyr	Lys	Trp	Thr	Ser 60
Pro	Pro	Gly	Val	Ile 65	Arg	Ile	Leu	Ser	Met 70	Leu	Ile	Ile	Val	Met 75
Cys	Ile	Ala	Ile	Phe 80	Ala	Cys	Val	Ala	Ser 85	Thr	Leu	Ala	Trp	Asp 90
Arg	Gly	Tyr	Gly	Thr 95	Ser	Leu	Leu	Gly	Gly 100	Ser	Val	Gly	Tyr	Pro 105
Tyr	Gly	Gly	Ser	Gly 110	Phe	Gly	Ser	Tyr	Gly 115	Ser	Gly	Tyr	Gly	Tyr 120
Gly	Tyr	Gly	Tyr	Gly 125	Tyr	Gly	Tyr	Gly	Gly 130	Tyr	Thr	Asp	Pro	Arg 135
Ala	Ala	Lys	Gly	Phe 140	Met	Leu	Ala	Met	Ala 145	Ala	Phe	Cys	Phe	Ile 150
Ala	Ala	Leu	Val	Ile 155	Phe	Val	Thr	Ser	Val 160	Ile	Arg	Ser	Glu	Met 165
Ser	Arg	Thr	Arg	Arg 170	Tyr	Tyr	Leu	Ser	Val 175	Ile	Ile	Val	Ser	Ala 180
Ile	Leu	Gly	Ile	Met 185	Val	Phe	Ile	Ala	Thr 190	Ile	Val	Tyr	Ile	Met 195
Gly	Val	Asn	Pro	Thr 200	Ala	Gln	Ser	Ser	Gly 205	Ser	Leu	Tyr	Gly	Ser 210
Gln	Ile	Tyr	Ala	Leu 215	Cys	Asn	Gln	Phe	Tyr 220	Thr	Pro	Ala	Ala	Thr 225
Gly	Leu	Tyr	Val	Asp 230	Gln	Tyr	Leu	Tyr	His 235	Tyr	Cys	Val	Val	Asp 240
Pro	Gln	Glu	Ala	Ile 245	Ala	Ile	Val	Leu	Gly 250	Phe	Met	Ile	Ile	Val 255

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
2	2	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
3	3	2	1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
4	4	3	2	1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
5	5	4	3	2	1	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				

[illegible]

7/7

Gly Tyr Gly Gly Tyr Thr Asp Pro Arg
20

(6) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: peptide

(v) FRAGMENT TYPE: synthetic peptide

(ix) FEATURE:

- (D) OTHER INFORMATION: construct used in experiments

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Asn	His	Tyr	Ala	Pro	Ser	Asn	Asp	Ile	Tyr	Gly	Gly	Glu	Met	Val
				5					10					15
His	Arg	Pro	Met	Leu										
				20										

(7) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: peptide

(v) FRAGMENT TYPE: synthetic peptide

(ix) FEATURE:

- (D) OTHER INFORMATION: construct used in experiments

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Ala	Ser	Gln	Gln	Val	Tyr	Arg	Lys	Asp	Pro	Cys
				5					10	